

RECEIVED

NOV 13 2001

TECH CENTER 1600/2900

#14/50
11-250
OIPE

RAW SEQUENCE LISTING

DATE: 09/20/2001

PATENT APPLICATION: US/09/489,198A

TIME: 12:57:07

Input Set : A:\06501-054001.txt

Output Set: N:\CRF3\09202001\I489198A.raw

PS

ENTERED

4 <110> APPLICANT: Kato, Shigeaki
 5 Takeyama, Ken-Ichi
 6 Kitanaka, Sachiko
 8 <120> TITLE OF INVENTION: GENE SCREENING METHOD USING NUCLEAR
 9 RECEPTOR
 11 <130> FILE REFERENCE: 06501-054001
 13 <140> CURRENT APPLICATION NUMBER: 09/489,198A
 14 <141> CURRENT FILING DATE: 2000-01-20
 16 <150> PRIOR APPLICATION NUMBER: PCT/JP98/03280
 17 <151> PRIOR FILING DATE: 1998-07-22
 19 <150> PRIOR APPLICATION NUMBER: JP 9/212624
 20 <151> PRIOR FILING DATE: 1997-07-22
 22 <160> NUMBER OF SEQ ID NOS: 4
 24 <170> SOFTWARE: FastSEQ for Windows Version 4.0
 26 <210> SEQ ID NO: 1
 27 <211> LENGTH: 507
 28 <212> TYPE: PRT
 29 <213> ORGANISM: Mus musculus
 31 <400> SEQUENCE: 1
 32 Met Thr Gln Ala Val Lys Leu Ala Ser Arg Val Phe His Arg Ile His
 33 1 5 10 15
 34 Leu Pro Leu Gln Leu Asp Ala Ser Leu Gly Ser Arg Gly Ser Glu Ser
 35 20 25 30
 36 Val Leu Arg Ser Leu Ser Asp Ile Pro Gly Pro Ser Thr Leu Ser Phe
 37 35 40 45
 38 Leu Ala Glu Leu Phe Cys Lys Gly Gly Leu Ser Arg Leu His Glu Leu
 39 50 55 60
 40 Gln Val His Gly Ala Ala Arg Tyr Gly Pro Ile Trp Ser Gly Ser Phe
 41 65 70 75 80
 42 Gly Thr Leu Arg Thr Val Tyr Val Ala Asp Pro Thr Leu Val Glu Gln
 43 85 90 95
 44 Leu Leu Arg Gln Glu Ser His Cys Pro Glu Arg Cys Ser Phe Ser Ser
 45 100 105 110
 46 Trp Ala Glu His Arg Arg Arg His Gln Arg Ala Cys Gly Leu Leu Thr
 47 115 120 125
 48 Ala Asp Gly Glu Glu Trp Gln Arg Leu Arg Ser Leu Leu Ala Pro Leu
 49 130 135 140
 50 Leu Leu Arg Pro Gln Ala Ala Gly Tyr Ala Gly Thr Leu Asp Asn
 51 145 150 155 160
 52 Val Val Arg Asp Leu Val Arg Arg Leu Arg Arg Gln Arg Gly Arg Gly
 53 165 170 175
 54 Ser Gly Leu Pro Gly Leu Val Leu Asp Val Ala Gly Glu Phe Tyr Lys
 55 180 185 190
 56 Phe Gly Leu Glu Ser Ile Gly Ala Val Leu Leu Gly Ser Arg Leu Gly
 57 195 200 205
 58 Cys Leu Glu Ala Glu Val Pro Pro Asp Thr Glu Thr Phe Ile His Ala
 59 210 215 220

RAW SEQUENCE LISTING

DATE: 09/20/2001

PATENT APPLICATION: US/09/489,198A

TIME: 12:57:07

Input Set : A:\06501-054001.txt

Output Set: N:\CRF3\09202001\I489198A.raw

```

60 Val Gly Ser Val Phe Val Ser Thr Leu Leu Thr Met Ala Met Pro Asn
61 225                230                235                240
62 Trp Leu His His Leu Ile Pro Gly Pro Trp Ala Arg Leu Cys Arg Asp
63                245                250                255
64 Trp Asp Gln Met Phe Ala Phe Ala Gln Arg His Val Glu Leu Arg Glu
65                260                265                270
66 Gly Glu Ala Ala Met Arg Asn Gln Gly Lys Pro Glu Glu Asp Met Pro
67                275                280                285
68 Ser Gly His His Leu Thr His Phe Leu Phe Arg Glu Lys Val Ser Val
69                290                295                300
70 Gln Ser Ile Val Gly Asn Val Thr Glu Leu Leu Ala Gly Val Asp
71 305                310                315                320
72 Thr Val Ser Asn Thr Leu Ser Trp Thr Leu Tyr Glu Leu Ser Arg His
73                325                330                335
74 Pro Asp Val Gln Thr Ala Leu His Ser Glu Ile Thr Ala Gly Thr Arg
75                340                345                350
76 Gly Ser Cys Ala His Pro His Gly Thr Ala Leu Ser Gln Leu Pro Leu
77                355                360                365
78 Leu Lys Ala Val Ile Lys Glu Val Leu Arg Leu Tyr Pro Val Val Pro
79                370                375                380
80 Gly Asn Ser Arg Val Pro Asp Arg Asp Ile Arg Val Gly Asn Tyr Val
81 385                390                395                400
82 Ile Pro Gln Asp Thr Leu Val Ser Leu Cys His Tyr Ala Thr Ser Arg
83                405                410                415
84 Asp Pro Thr Gln Phe Pro Asp Pro Asn Ser Phe Asn Pro Ala Arg Trp
85                420                425                430
86 Leu Gly Glu Gly Pro Thr Pro His Pro Phe Ala Ser Leu Pro Phe Gly
87                435                440                445
88 Phe Gly Lys Arg Ser Cys Ile Gly Arg Arg Leu Ala Glu Leu Glu Leu
89                450                455                460
90 Gln Met Ala Leu Ser Gln Ile Leu Thr His Phe Glu Val Leu Pro Glu
91 465                470                475                480
92 Pro Gly Ala Leu Pro Ile Lys Pro Met Thr Arg Thr Val Leu Val Pro
93                485                490                495
94 Glu Arg Ser Ile Asn Leu Gln Phe Val Asp Arg
95                500                505
97 <210> SEQ ID NO: 2
98 <211> LENGTH: 508
99 <212> TYPE: PRT
100 <213> ORGANISM: Homo sapiens
102 <400> SEQUENCE: 2
103 Met Thr Gln Thr Leu Lys Tyr Ala Ser Arg Val Phe His Arg Val Arg
104 1                5                10                15
105 Trp Ala Pro Glu Leu Gly Ala Ser Leu Gly Tyr Arg Glu Tyr His Ser
106                20                25                30
107 Ala Arg Arg Ser Leu Ala Asp Ile Pro Gly Pro Ser Thr Pro Ser Phe
108                35                40                45
109 Leu Ala Glu Leu Phe Cys Lys Gly Gly Leu Ser Arg Leu His Glu Leu
110                50                55                60

```

RAW SEQUENCE LISTING

DATE: 09/20/2001

PATENT APPLICATION: US/09/489,198A

TIME: 12:57:07

Input Set : A:\06501-054001.txt

Output Set: N:\CRF3\09202001\I489198A.raw

```

111 Gln Val Gln Gly Ala Ala His Phe Gly Pro Val Trp Leu Ala Ser Phe
112 65 70 75 80
113 Gly Thr Val Arg Thr Val Tyr Val Ala Ala Pro Ala Leu Val Glu Glu
114 85 90 95
115 Leu Leu Arg Gln Glu Gly Pro Arg Pro Glu Arg Cys Ser Phe Ser Pro
116 100 105 110
117 Trp Thr Glu His Arg Arg Cys Arg Gln Arg Ala Cys Gly Leu Leu Thr
118 115 120 125
119 Ala Glu Gly Glu Glu Trp Gln Arg Leu Arg Ser Leu Leu Ala Pro Leu
120 130 135 140
121 Leu Leu Arg Pro Gln Ala Ala Ala Arg Tyr Ala Gly Thr Leu Asn Asn
122 145 150 155 160
123 Val Val Cys Asp Leu Val Arg Arg Leu Arg Arg Gln Arg Gly Arg Gly
124 165 170 175
125 Thr Gly Pro Pro Ala Leu Val Arg Asp Val Ala Gly Glu Phe Tyr Lys
126 180 185 190
127 Phe Gly Leu Glu Gly Ile Ala Ala Val Leu Leu Gly Ser Arg Leu Gly
128 195 200 205
129 Cys Leu Glu Ala Gln Val Pro Pro Asp Thr Glu Thr Phe Ile Arg Ala
130 210 215 220
131 Val Gly Ser Val Phe Val Ser Thr Leu Leu Thr Met Ala Met Pro His
132 225 230 235 240
133 Trp Leu Arg His Leu Val Pro Gly Pro Trp Gly Arg Leu Cys Arg Asp
134 245 250 255
135 Trp Asp Gln Met Phe Ala Phe Ala Gln Arg His Val Glu Arg Arg Glu
136 260 265 270
137 Ala Glu Ala Ala Met Arg Asn Gly Gly Gln Pro Glu Lys Asp Leu Glu
138 275 280 285
139 Ser Gly Ala His Leu Thr His Phe Leu Phe Arg Glu Glu Leu Pro Ala
140 290 295 300
141 Gln Ser Ile Leu Gly Asn Val Thr Glu Leu Leu Ala Gly Val Asp
142 305 310 315 320
143 Thr Val Ser Asn Thr Leu Ser Trp Ala Leu Tyr Glu Leu Ser Arg His
144 325 330 335
145 Pro Glu Val Gln Thr Ala Leu His Ser Glu Ile Thr Ala Ala Leu Ser
146 340 345 350
147 Pro Gly Ser Ser Ala Tyr Pro Ser Ala Thr Val Leu Ser Gln Leu Pro
148 355 360 365
149 Leu Leu Lys Ala Val Val Lys Glu Val Leu Arg Leu Tyr Pro Val Val
150 370 375 380
151 Pro Gly Asn Ser Arg Val Pro Asp Lys Asp Ile His Val Gly Asp Tyr
152 385 390 395 400
153 Ile Ile Pro Lys Asn Thr Leu Val Thr Leu Cys His Tyr Ala Thr Ser
154 405 410 415
155 Arg Asp Pro Ala Gln Phe Pro Glu Pro Asn Ser Phe Arg Pro Ala Arg
156 420 425 430
157 Trp Leu Gly Glu Gly Pro Thr Pro His Pro Phe Ala Ser Leu Pro Phe
158 435 440 445
159 Gly Phe Gly Lys Arg Ser Cys Met Gly Arg Arg Leu Ala Glu Leu Glu

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/489,198A

DATE: 09/20/2001

TIME: 12:57:07

Input Set : A:\06501-054001.txt

Output Set: N:\CRF3\09202001\I489198A.raw

```

160      450      455      460
161 Leu Gln Met Ala Leu Ala Gln Ile Leu Thr His Phe Glu Val Gln Pro
162 465      470      475      480
163 Glu Pro Gly Ala Ala Pro Val Arg Pro Lys Thr Arg Thr Val Leu Val
164      485      490      495
165 Pro Glu Arg Ser Ile Asn Leu Gln Phe Leu Asp Arg
166      500      505
168 <210> SEQ ID NO: 3
169 <211> LENGTH: 2386
170 <212> TYPE: DNA
171 <213> ORGANISM: Mus musculus
173 <220> FEATURE:
174 <221> NAME/KEY: CDS
175 <222> LOCATION: (30)...(1550)
177 <400> SEQUENCE: 3
178 ctctcgaagc agactcccca aacacagac atg acc cag gca gtc aag ctc gcc      53
179      Met Thr Gln Ala Val Lys Leu Ala
180      1      5
182 tcc aga gtt ttt cac cga atc cac ctg cct ctg cag ctg gat gcc tcg      101
183 Ser Arg Val Phe His Arg Ile His Leu Pro Leu Gln Leu Asp Ala Ser
184 10      15      20
186 ctg ggc tcc aga ggc agt gag tcg gtt ctc cgg agc ttg tct gac atc      149
187 Leu Gly Ser Arg Gly Ser Glu Ser Val Leu Arg Ser Leu Ser Asp Ile
188 25      30      35      40
190 cct ggg ccc tct aca ctc agc ttc ctg gct gaa ctc ttc tgc aaa ggg      197
191 Pro Gly Pro Ser Thr Leu Ser Phe Leu Ala Glu Leu Phe Cys Lys Gly
192      45      50      55
194 ggg ctg tcc agg ctg cat gaa ctg cag gtg cat ggc gct gcg cgg tac      245
195 Gly Leu Ser Arg Leu His Glu Leu Gln Val His Gly Ala Ala Arg Tyr
196      60      65      70
198 ggg cca ata tgg tct ggc agc ttt ggg aca ctt cgc aca gtt tac gtt      293
199 Gly Pro Ile Trp Ser Gly Ser Phe Gly Thr Leu Arg Thr Val Tyr Val
200      75      80      85
202 gcc gac cct aca ctt gtg gag cag ctc ctg cga caa gaa agt cac tgt      341
203 Ala Asp Pro Thr Leu Val Glu Gln Leu Leu Arg Gln Glu Ser His Cys
204      90      95      100
206 cca gag cgc tgt agt ttc tca tca tgg gca gag cac cgt cgc cgc cac      389
207 Pro Glu Arg Cys Ser Phe Ser Ser Trp Ala Glu His Arg Arg Arg His
208 105      110      115      120
210 cag cgt gct tgc gga ttg cta acg gcg gat ggt gaa gaa tgg cag agg      437
211 Gln Arg Ala Cys Gly Leu Leu Thr Ala Asp Gly Glu Glu Trp Gln Arg
212      125      130      135
214 ctc cga agt ctt ctg gcc ccg ctc ctc ctc cgg cca caa gca gcc gcg      485
215 Leu Arg Ser Leu Leu Ala Pro Leu Leu Leu Arg Pro Gln Ala Ala Ala
216      140      145      150
218 ggc tat gct gga act ctg gac aac gtg gtc cgt gac ctt gtg cga cga      533
219 Gly Tyr Ala Gly Thr Leu Asp Asn Val Val Arg Asp Leu Val Arg Arg
220      155      160      165
222 cta agg cgc cag cgg gga cgt ggc tct ggg cta ccc ggc cta gtt ctg      581

```

RAW SEQUENCE LISTING

DATE: 09/20/2001

PATENT APPLICATION: US/09/489,198A

TIME: 12:57:07

Input Set : A:\06501-054001.txt

Output Set: N:\CRF3\09202001\I489198A.raw

223	Leu	Arg	Arg	Gln	Arg	Gly	Arg	Gly	Ser	Gly	Leu	Pro	Gly	Leu	Val	Leu	
224		170					175					180					
226	gac	gtg	gca	gga	gag	ttt	tac	aaa	ttt	ggc	cta	gaa	agt	ata	ggc	gcg	629
227	Asp	Val	Ala	Gly	Glu	Phe	Tyr	Lys	Phe	Gly	Leu	Glu	Ser	Ile	Gly	Ala	
228	185					190					195					200	
230	gtg	ctg	ctg	gga	tgc	cgc	ctg	ggc	tgc	cta	gag	gct	gaa	gtc	cct	cct	677
231	Val	Leu	Leu	Gly	Ser	Arg	Leu	Gly	Cys	Leu	Glu	Ala	Glu	Val	Pro	Pro	
232					205					210					215		
234	gac	aca	gaa	acc	ttc	ata	cat	gca	gtg	ggc	tca	gtg	ttt	gtg	tct	aca	725
235	Asp	Thr	Glu	Thr	Phe	Ile	His	Ala	Val	Gly	Ser	Val	Phe	Val	Ser	Thr	
236				220						225				230			
238	ctc	ttg	acc	atg	gcg	atg	ccc	aac	tgg	ttg	cac	cac	ctt	ata	cct	gga	773
239	Leu	Leu	Thr	Met	Ala	Met	Pro	Asn	Trp	Leu	His	His	Leu	Ile	Pro	Gly	
240			235					240					245				
242	ccc	tgg	gcc	cgc	ctc	tgc	cga	gac	tgg	gat	cag	atg	ttt	gcc	ttt	gcc	821
243	Pro	Trp	Ala	Arg	Leu	Cys	Arg	Asp	Trp	Asp	Gln	Met	Phe	Ala	Phe	Ala	
244		250				255					260						
246	cag	agg	cac	gtg	gag	ctg	cga	gaa	ggt	gaa	gct	gcg	atg	agg	aac	cag	869
247	Gln	Arg	His	Val	Glu	Leu	Arg	Glu	Gly	Glu	Ala	Ala	Met	Arg	Asn	Gln	
248	265				270					275						280	
250	gga	aag	cct	gag	gag	gat	atg	ccg	tct	ggg	cat	cac	tta	acc	cac	ttc	917
251	Gly	Lys	Pro	Glu	Glu	Asp	Met	Pro	Ser	Gly	His	His	Leu	Thr	His	Phe	
252				285						290					295		
254	ctt	ttt	cgg	gaa	aag	gtg	tct	gtc	cag	tcc	ata	gtg	ggg	aat	gtg	aca	965
255	Leu	Phe	Arg	Glu	Lys	Val	Ser	Val	Gln	Ser	Ile	Val	Gly	Asn	Val	Thr	
256			300						305				310				
258	gag	cta	cta	ctg	gct	gga	gtg	gac	acg	gta	tcc	aat	acg	ctc	tcc	tgg	1013
259	Glu	Leu	Leu	Leu	Ala	Gly	Val	Asp	Thr	Val	Ser	Asn	Thr	Leu	Ser	Trp	
260			315					320					325				
262	aca	ctc	tat	gag	ctt	tcc	cgg	cac	ccc	gat	gtc	cag	act	gca	ctc	cac	1061
263	Thr	Leu	Tyr	Glu	Leu	Ser	Arg	His	Pro	Asp	Val	Gln	Thr	Ala	Leu	His	
264		330				335					340						
266	tct	gag	atc	aca	gct	ggg	acc	cgt	ggc	tcc	tgt	gcc	cac	ccc	cat	ggc	1109
267	Ser	Glu	Ile	Thr	Ala	Gly	Thr	Arg	Gly	Ser	Cys	Ala	His	Pro	His	Gly	
268	345				350					355						360	
270	act	gct	ctg	tcc	cag	ctg	ccc	ctg	tta	aag	gct	gtg	atc	aaa	gaa	gtg	1157
271	Thr	Ala	Leu	Ser	Gln	Leu	Pro	Leu	Leu	Lys	Ala	Val	Ile	Lys	Glu	Val	
272				365					370						375		
274	ttg	aga	ttg	tac	cct	gtg	gta	cct	ggg	aat	tcc	cgt	gtc	cca	gac	aga	1205
275	Leu	Arg	Leu	Tyr	Pro	Val	Val	Pro	Gly	Asn	Ser	Arg	Val	Pro	Asp	Arg	
276			380						385				390				
278	gac	atc	cgt	gta	gga	aac	tat	gta	att	ccc	caa	gat	acg	cta	gtc	tcc	1253
279	Asp	Ile	Arg	Val	Gly	Asn	Tyr	Val	Ile	Pro	Gln	Asp	Thr	Leu	Val	Ser	
280			395				400						405				
282	cta	tgt	cac	tat	gcc	act	tca	agg	gac	ccc	aca	cag	ttt	cca	gac	ccc	1301
283	Leu	Cys	His	Tyr	Ala	Thr	Ser	Arg	Asp	Pro	Thr	Gln	Phe	Pro	Asp	Pro	
284		410				415					420						
286	aac	tct	ttt	aat	cca	gct	cgc	tgg	ctg	ggg	gag	ggt	ccg	acc	ccc	cac	1349
287	Asn	Ser	Phe	Asn	Pro	Ala	Arg	Trp	Leu	Gly	Glu	Gly	Pro	Thr	Pro	His	

→ Use of n and/or Xaa has been detected in the Sequence Listing.
 Review the Sequence Listing to insure a corresponding
 explanation is presented in the <220> to <223> fields of
 each sequence using n or Xaa.

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/489,198A

DATE: 09/20/2001

TIME: 12:57:08

Input Set : A:\06501-054001.txt

Output Set: N:\CRF3\09202001\I489198A.raw

L:473 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4